

UNIVERSITY OF COPENHAGEN



Kausal inferens i statskundskaben

Hariri, Jacob Gerner

Published in:
Politica

Publication date:
2012

Citation for published version (APA):
Hariri, J. G. (2012). Kausal inferens i statskundskaben. *Politica*, 44(2), 184-201.

Jacob Gerner Hariri

Kausal inferens i statskundskaben¹

Kausal inferens er ikke blot tidens mode i empirisk samfundsforskning, det er også en forudsætning for at kunne designe eller evaluere policy-interventioner i praksis. Denne artikel definerer den kausale relation og diskuterer i forlængelse deraf, hvilke kriterier der må stilles, før det er identificeret, hvorvidt og i hvilken grad et fænomen er årsag til et andet. Artiklen viser, hvorfor det formelle krav her til må være, at den uafhængige variabel er eksogen, og der argumenteres for, at de tre kriterier tidsrækkefølge, fravær af spuriøsitet og teoretisk begrundelse kan være upræcise i samfundsfaglige observationsstudier. Desuden illustrerer et eksempel, hvordan samfundsfaglige observationsstudier kan søge at leve op til det formelle kriterium for kausal inferens.

Kausal inferens – hvornår og i hvilken grad et fænomen kan siges at være årsag til et andet – er et af de vanskeligste og mest kontroversielle metodiske og videnskabsteoretiske emner. Det er vanskeligt, fordi det kan siges at involvere noget ikke observerbart, nemlig den kontrafaktiske situation, hvori årsagen ikke var indtruffet. Det er kontroversielt, fordi kausalitet igennem lang tid af en positivistisk forskningstradition blev afvist som grundlæggende uvidenskabeligt: Verden bestod af korrelationer, kausal inferens var *story telling*.

Denne artikel diskuterer kausal inferens i statskundskaben. Det er i dag velkendt og accepteret, at kausal inferens forudsætter korrelation, men at korrelation omvendt ikke er tilstrækkeligt for kausal inferens. Derfor opstilles normalt et sæt af kriterier, som en observeret korrelation skal opfylde, før den kan betragtes som kausal. I statskundskaben er disse kriterier (Andersen, 2010: 104):

1. Tidsrækkefølge: Den uafhængige variabel indtræffer før den afhængige.
2. Teori: Den kausale sammenhæng skal være teoretisk begrundet.
3. Fravær af spuriøsitet: Der må ikke være andre variable, der påvirker både den afhængige og den uafhængige variabel.

Der argumenteres i denne artikel for, at det første kriterium ikke hjælper til at identificere en kausal effekt, hvis der er stiafhængighed i det empiriske system, der undersøges. Det andet kriterium er ikke nødvendigt for at identificere en kausal effekt, og det hjælper ikke til at udsondre den kausale effekt fra den observerede korrelation. Desuden argumenteres der for, at det tredje kriterium

skal skærpes: Der må ikke være andre variable, som *har sammenhæng med* både den uafhængige variabel og den afhængige variabel. Artiklen er således ikke et katalog over empiriske metoder til kausal identifikation men et forsøg på mere grundlæggende at afdække og diskutere de bagvedliggende kriterier herfor.²

En kausal relation defineres som en asymmetrisk relation mellem to fænomener, årsag og virkning, hvor den betingede sandsynlighed for virkning, når årsag er til stede, adskiller sig fra den ubetingede sandsynlighed for virkning. I overensstemmelse med denne definition vises det at for at kunne identificere den kausale virkning af en uafhængig variabel, skal denne være eksogen, hvor eksogenitet betyder, at (a) den afhængige variabel ikke virker tilbage på den uafhængige (ingen omvendt kausalitet) samt, at (b) ingen af de faktorer, der korrelerer med den uafhængige variabel, må korrelere med de faktorer, der korrelerer med den afhængige.

Eksogenitet er ikke et absolut krav i en praktisk empirisk undersøgelse – lige så lidt som perfekt normalfordelte fejllid i praksis er et absolut krav i en regressionsanalyse. Men det er artiklens påstand, at det er en præcis formulering af den formelle betingelse for kausal identifikation, og at en præcis formulering af et problem er det bedste udgangspunkt for at tage hånd om det. I praktisk arbejde må man så forsøge, så godt det lader sig gøre, at leve op til disse betingelser. I et afsluttende afsnit illustreres det derfor med et praktisk eksempel, hvordan observationsstudier kan designes, så den uafhængige variabel kommer nærmere at opfylde kravet om eksogenitet.

Først opridses dog et par linjer af kausalitetsbegrebets historie, ligesom der afsættes tid til at afklare og definere den kausale relation. Selvom artiklens ærinde er epistemologisk, må vi definere den kausale relation, før vi kan identificere den.

Et rids af kausalitetens historie

Tænkningen om kausalitet og kausal inferens har dybe videnskabshistoriske rødder. Således sondrede Aristoteles fx mellem fire årsagstyper, men betonedes selv den såkaldt *finale* årsag – fænomenernes naturlige tilstand, der drager dem imod deres naturlige form. Den finale årsag til, at en sten falder, er, at stenens naturlige form findes på jorden. Damp derimod stiger til vejrs, fordi dampens naturlige form er luftig (Sløk et al., 2007: 94-95). I overensstemmelse med dette kausalitetsbegreb var videnskabens rolle i antikken at begribe hensigten og den naturlige form bag ting og fænomener og klassificere dem derefter (Sløk et al., 2007: 226f.).

Med fremkomsten af mekanikken og den moderne fysik i renæssancen ændredes både kausalitetsbegrebet og synet på videnskaben. Videnskabens rolle

blev nu ikke længere set som klassifikatorisk men mere dynamisk at begribe fænomenernes opståen og deres indbyrdes relationer. Årsagen til et fænomen var nu ikke længere dets essens, men de andre fænomener, der har medført det. Kausalitetsbegrebet blev hermed den virkende årsag – den som frembringer sin virkning.

I samfundsvidenskaben i dag er det dominerende kausalitetsbegreb også den virkende årsag. Spørgsmålet om, hvordan vi identificerer, hvorvidt og hvor meget årsag påvirker virkning (altså spørgsmålet om kausal inferens), har været omstridt i videnskaberne i hvert fald siden David Humes intellektuelle hovedrengøring, der henviste sofisteri og snik-snak til bålet: "*Does it contain any abstract reasoning concerning quantity and number?* No. *Does it contain any experimental reasoning concerning matter of fact and existence?* No. Commit it then to the flames: for it can contain nothing but sophistry and illusion" (citeret i Hoover, 2001: 10)

Hume forsøgte at formulere en empirisk verificerbar tilgang til kausalitet. Han afviste derfor, at kausalitet kunne bestå i en nødvendig relation mellem årsag og virkning, da nødvendighed ikke er empirisk observerbar. I stedet definerede Hume en kausal relation ved hjælp af de tre observerbare kriterier, at (i) årsag skal indtræde før virkning, at (ii) årsag og virkning skal være rumligt eller tidsligt tilstødende fænomener, samt at (iii) årsag og virkning altid optræder sammen. Humes tre kriterier blev siden kritiseret for ikke at kunne skelne sammenhænge fra årsagssammenhænge. Således blev det allerede i samtiden påpeget, at ifølge Humes kriterier ville mandag aften være årsag til tirsdag morgen (Psillos, 2009: 136).

Inspireret af Humes afvisning af det ikke-observerbare, afviste positivister langt op i det 20. århundrede kausalitet som uvidenskabeligt (Keuzenkamp, 2000). Videnskabens rolle var at observere og registrere tilbagevendende sammenhænge: Vi kan observere to fænomener og måle deres samvariation, men at kalde det ene årsag og det andet virkning er *story telling*, ikke videnskab. "Science is measurement", kausalitet er snik-snak (Keuzenkamp 2000: viii). Det var som eksponent for denne holdning, at Bertrand Russell i 1913 sammenlignede kausalitet med monarkiet: et levn fra fortiden, der kun overlever, fordi det fejlagtigt regnes for harmløst (Russell, 1913).

Igennem de seneste årtier har denne radikale positivisme fortaget sig. Nu er det ikke blot videnskabeligt accepteret men højeste mode at søge at identificere kausale sammenhænge. Kausalitet og kausal inferens er "back with a vengeance", som videnskabsteoretikeren Nancy Cartwright i 2007 formulerede det (2007: 1). Efter denne forfatters mening bidrager genoplivningen af den kausale inferens til at højne statskundskabens praktiske relevans. Når *statskund-*

skaberen konstruerer eller evaluerer et politisk tiltag, er det ikke nok som de radikale positivister at betragte verden som et sæt af relationer, hvor fænomener samvarierer. Det er ikke nok for eksempel at konstatere, at skolebørn, der spiser frokost, klarer sig bedre end børn, der ikke gør, eller at deltagere i jobtræningskurser kommer hurtigere i arbejde end ledige, der ikke er i jobtræning. Det relevante (men uhyre vanskelige) spørgsmål er, hvorvidt frokosten eller jobtræningen i sig selv er kausalt relevant for karakterniveauet henholdsvis beskæftigelsesraten. For at kunne intervenere på fornuftig vis i samfundet, må vi vide, hvilken virkning den pågældende intervention har. Derfor er en empirisk relevant statskundskab også en statskundskab, der søger at afdække årsagssammenhænge og således eksplicit bestræber sig på at opnå kausal identifikation.

Dette er imidlertid vanskeligere end som så. Som nævnt indledningsvis er der tradition for ligesom Hume at opstille kriterier, som en observeret sammenhæng skal opfylde, før den kan udlægges som en årsagssammenhæng. I statskundskaben anvendes traditionelt de tre ovennævnte kriterier. For så vidt som disse kriterier er nøglen til kausal inferens, er det naturligvis væsentligt, at de diskuteres. Først vender vi os dog mod det begrebslige og forsøger at afklare, hvad der menes med en kausal relation.

Begrebsafklaring: Hvad er en kausal relation?

Når vi i daglig tale siger, at x forårsager y , understår vi ikke dermed, at y virker tilbage på x (Mackie, 1974; Simon, 1953; Winship og Sobel, 2001). Hvis vi betragter de to udsagn (i) ”vejret påvirker mit valg af overtøj” og (ii) ”jobtræning påvirker lediges sandsynlighed for at komme i beskæftigelse”, er det i udsagn (i) oplagt, at y ikke virker tilbage på x – mit tøjvalg påvirker ikke vejret. Derimod er det i udsagn (ii) formentlig tilfældet, at y virker tilbage på x – hvorvidt og hvor længe man deltager i jobtræning, afhænger af sandsynligheden for at finde arbejde. Pointen er her, at denne empiriske feedback-effekt ikke er indeholdt i udsagn (ii) og heller ikke i selve kausalitetsbegrebet: Uanset om virkeligheden er reciprok (således at y i et givent tilfælde også påvirker x), er kausalitetsbegrebet asymmetrisk. Årsag forholder sig til virkning på en måde, som virkning ikke kan siges at forholde sig til årsag. Årsag forårsager virkning, det modsatte er ikke tilfældet. En årsagssammenhæng er derfor en asymmetrisk relation mellem to fænomener (årsag og virkning), hvori årsag kan siges at have kausal prioritet (Mackie, 1974: 160).

Dette er desværre mindre trivielt, end det lyder. For uanset om årsagssammenhængen betragtes probabilistisk (således at årsag forandrer sandsynligheden for virkning) eller deterministisk (således at årsag er nødvendig og/eller

tilstrækkelig for virkning), er relationen logisk set *ikke* asymmetrisk. Dette er problematisk, når nu kausalitetsbegrebet er det.

Vi benytter her kausalitetsbegrebet probabilistisk. Det er i overensstemmelse med daglig brug og med meget empirisk arbejde i samfundsvidenskaberne (Suppes, 1970; Mahoney, 2008). Hvis vi i daglig tale siger, "du får hovedpine, hvis du drikker mere vin", mener vi dermed, at sandsynligheden for at vågne med hovedpine er større, hvis du fortsætter med at drikke, end hvis du ikke fortsætter. Vi mener ikke, at du med garanti undgår hovedpine, hvis du stopper drikkeriet (hvorved "mere vin" ville være en nødvendig årsag til hovedpine), og vi mener heller ikke, at hovedpinen med sikkerhed indtræffer, hvis drikkeriet fortsætter (hvorved "mere vin" ville være en tilstrækkelig årsag til hovedpine).

Vi siger altså, at x har en kausal effekt på y , hvis den betingede sandsynlighed for y givet x er større (eller mindre) end den ubetingede sandsynlighed for y : $P(y|x) \neq P(y)$. Vi forstår dermed årsagen som en "sandsynligheds-forandrer" (Mahoney, 2008). Denne forståelse er praktisk anvendelig, fordi den er bred nok til at omfatte både nødvendige årsager, tilstrækkelige årsager, nødvendige og tilstrækkelige årsager samt såkaldte *inus*-årsager.³ Den er imidlertid tilsvarende upræcis, fordi den ikke formår at skelne mellem disse. Ved at observere en forandret betinget sandsynlighed for y , kan det ikke umiddelbart afgøres, om relationen mellem x og y har karakter af nødvendighed eller tilstrækkelighed.

Værre er det dog, at sammenhængen ikke er asymmetrisk. Hvis den betingede sandsynlighed for y givet x er større end den ubetingede sandsynlighed for y , så gælder det modsatte også – at den betingede sandsynlighed for x givet y er større end den ubetingede sandsynlighed for x .⁴ Eller mere mundret: Hvis x er årsag til y (i betydningen sandsynligheds-forandrer), er y også årsag til x . Samme problem opstår for de deterministiske opfattelser af kausalitet. For eksempel gælder det, at hvis en årsag er nødvendig og tilstrækkelig for en virkning, følger det logisk, at virkningen også er nødvendig og tilstrækkelig for årsagen: Vi kan aldrig observere y uden x , ligesom vi aldrig kan observere x uden y (Mackie, 1974: 160ff).

Det viser sig altså, at både en deterministisk og meget stram fortolkning af årsag som en nødvendig og tilstrækkelig betingelse samt en probabilistisk og mere løs fortolkning af årsag som sandsynligheds-forandrer løber ind i samme problem: Hvis årsag forårsager virkning, forårsager virkning årsag. Det er derfor ikke fyldestgørende simpelthen at definere kausalitet ved, at årsag forandrer sandsynligheden for virkning. I tillæg må vi eksplicit kræve, at relationen er asymmetrisk, således at årsag har kausal prioritet.

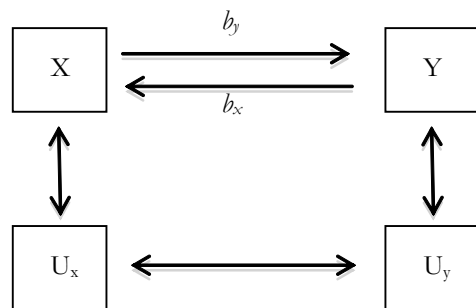
Efter disse overvejelser kan en kausal relation defineres som en *relation mellem årsag og virkning, som er asymmetrisk, fordi årsag har kausal prioritet, og hvor den betingede sandsynlighed for virkning givet årsag adskiller sig fra den ubetingede sandsynlighed for virkning*.

Denne definition er todelt: For det første skal der være en sammenhæng, og for det andet skal sammenhængen være asymmetrisk. Det første er let nok: Hvis der er korrelation mellem to fænomener, er den betingede sandsynlighed for det ene givet det andet anderledes end den ubetingede sandsynlighed for det ene. Det vanskelige punkt i definitionen er det andet led – for hvordan ved vi, om årsagen har kausal prioritet?

Næste afsnit viser, at for at kunne identificere størrelsen af en asymmetrisk årsagssammenhæng skal den uafhængige variabel være eksogen, hvor eksogenitet er defineret ved de to præcise betingelser (a) og (b) nævnt indledningsvis.

Eksogenitet som betingelse for kausal identifikation

Vi begynder fra det forsigtige (men for samfundsvidenskaben realistiske) udgangspunkt, at vi har observeret en sammenhæng mellem to fænomener målt ved de to variable, x og y , men at vi ikke ved, om x er årsag til y eller omvendt. Desuden er der en række andre faktorer (u_x), som varierer med x , ligesom der også er faktorer (u_y), der varierer med y . Figur 1 illustrerer denne situation, hvor alt i udgangspunktet hænger sammen (dog for visse faktors vedkommende



kun indirekte).

b_y angiver den kausale påvirkning fra x til y og er den størrelse, vi søger at identificere. Tilsvarende angiver b_x størrelsen af påvirkningen fra y til x . Situationen i Figur 1 er i overensstemmelse med forståelsen af årsag som sandsynlighedsforandrer, idet sandsynligheden for y er anderledes, når x er til stede, end når

x ikke er. Det er dog oplagt, at dette ikke nødvendigvis kan tilskrives x . Den betingede sandsynlighed for y givet x kan ændre sig, fordi y påvirker x , eller fordi u_x og u_y over indflydelse på x og y .

Figuren illustrerer dermed, at hvis vi skal identificere b_y – størrelsen af den kausale påvirkning fra x til y – skal vi stille betingelser, som (a) eliminerer pilen fra y til x samt (b) eliminerer pilen mellem u_x og u_y . For at se dette mere tydeligt, opskrives situationen som et system af to ligninger (Simon, 1953; 1954), hvor årsager står til højre for lighedstegnet, og virkninger til venstre (se Hoover, 2001: 38f.):

$$(1) \quad y = b_y x + a_y u_y$$

$$(2) \quad x = b_x y + a_x u_x$$

Ligning (1) viser y som funktion af x og "alt andet", mens ligning (2) viser x som funktion af y og "alt andet". Variablene x og y , der måler de to fænomener, vi interesserer os for, kan direkte observeres. Der er dog stadig fire ubekendte i systemet: Effekten af x på y (b_y), effekten af y på x (b_x), korrelationen mellem y og "alt andet" (a_y) samt korrelationen mellem x og "alt andet" (a_x). Vi kan alene ved at tælle ligninger og ubekendte indse, at dette system ikke kan identificere effekten af x på y (altså b_y): Til de fire ubekendte er der kun to ligninger. Derfor er vi nødt til at lægge yderligere to restriktioner ned over systemet, så antallet af ligninger svarer til antallet af ubekendte. De påkrævede restriktioner er kriterierne (a) og (b):

$$(a) \quad b_x = 0$$

$$(b) \quad E[u_x u_y] = 0$$

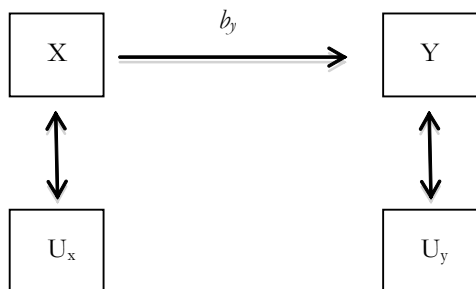
Det første af disse kriterier, $b_x = 0$, siger, at y ikke må påvirke x ; der er ikke omvendt kausalitet. Det andet, $E[u_x u_y] = 0$, siger at der ikke må være nogen korrelation mellem "alt det, som korrelerer med y " og "alt det, som korrelerer med x ". Det er altså ikke tilstrækkeligt, at den uafhængige og den afhængige variabel ikke påvirkes af samme tredje-variabel. Hvis blot x og y korrelerer med forskellige tredje-variable, som er indbyrdes forbundne, er (b) ikke længere opfyldt. Disse betingelser muliggør kausal identifikation, fordi det oprindelige ligningssystem nu kan omskrives til:

$$(1') \quad y = b_y x + a_y u_y$$

$$(2') \quad x = a_x u_x$$

Systemet med (1') og (2') har nu den ønskede kausale struktur, hvor x er (inus)-årsag til y , og b_y angiver størrelsen af den kausale påvirkning. x bestemmes i relation (2'), hvori y ikke indgår. Først derefter bestemmes y i relation (1'), hvori x og u_y indgår. x bestemmes dermed først og har kausal prioritet over y . Sammenhængen mellem x og y er inferentielt asymmetrisk, således forstået at vi kan slutte os fra tilstedeværelsen af x til (en del af) y , mens vi omvendt ikke kan slutte os fra y til x . Vi siger, at x er eksogen, fordi x bestemmes uden for det system, hvori y bestemmes. Og det er netop denne eksogenitet, der gør os i stand til at identificere den kausale effekt af x på y som b_y . Grafisk kan det nye ligningssystem illustreres som vist nedenfor.

Figur 2: Kausal identifikation af størrelsen af påvirkningen af x på y



Det er værd at bemærke, at b_y svarer præcist til OLS-estimatoren i en lineær regression.⁵ Dette understreger den praktiske relevans af vores overvejelser: Kun hvis (a) og (b) er opfyldt, kan OLS-estimatoren fortolkes kausalt. Hvis ikke er den blot et udtryk for korrelation.

Det skal understreges, at både (a) og (b) skal være opfyldt, før vi kan identificere en kausal påvirkning af x på y . Hvis (a) ikke er opfyldt, kan pilen fra y til x i Figur 1 ikke elimineres. Det betyder, at b_y ikke kan fortolkes som et kausalt estimat, idet b_y også indeholder feedback-effekten fra y på x . Hvis omvendt (b) ikke er opfyldt, kan pilen mellem u_x og u_y ikke elimineres. I dette tilfælde er en del af korrelationen mellem x og y spuriøs, selvom ingen af de to variable i sig selv korrelerer med samme tredje faktor. Dette bliver "fejltagtigt" opfanget af b_y , som derfor i denne situation ikke længere kan udlægges som den kausale effekt af x på y . James Mahoney (2008: 423) skriver, at det som regel er tilstrækkeligt, at (a) er opfyldt, men at kausale slutninger "strengt taget" forudsætter både (a) og (b). Denne formulering er uheldig. Hvis (b) ikke er opfyldt, ved vi ikke om estimatet af b_y i virkeligheden opfanger "noget af alt det andet, som korrelerer med både x og y ". Derfor er begge betingelser (a) og (b) nødvendige. De er

samtidig tilstrækkelige: Hvis der er korrelation, og den uafhængige variabel opfylder (a) og (b), udtrykker b_y den kausale påvirkning fra x til y .

Betingelse (b) er en skærpelse af formuleringen om, at samme variabel ikke må påvirke både den uafhængige og den afhængige variabel. I figur 1 er der korrelation mellem x og u_y (via u_x), men det er formelt set ikke korrekt, at u_y påvirker både x og y : Sammenhængen mellem den uafhængige variabel og en relevant tredjevariabel går her via en fjerde variabel, og det er tilstrækkeligt til at umuliggøre identifikation af den kausale påvirkning fra x på y . Det er i denne skærpede form, der i praktisk empirisk arbejde bør argumenteres for fravær af spuriøsitet.

Vi har nu defineret den kausale relation og demonstreret, at den uafhængige variabel skal opfylde (a) og (b) og i denne betydning være eksogen, før størrelsen af den kausale påvirkning fra den uafhængige på den afhængige variabel kan identificeres. Vi fortsætter med at diskutere, hvorvidt kriteriet om tidsrækkefølge og teoretisk begrundelse også kan siges at identificere en kausal relation.

Tidsrækkefølge og kausal identifikation

Betingelsen om, at "værdien på den uafhængige variabel bliver bestemt før værdien på den afhængige variabel" (Andersen, 2010: 104), stemmer overens med vores dagligdags oplevelse af verden: Billardkuglen bevæger sig først *efter*, vi har stødt til den. Desuden introducerer betingelsen som krævet en asymmetri i relationen mellem årsag og virkning: Årsag er nu det fænomen, som kommer før virkning, og i hvis tilstedeværelse den betingede sandsynlighed for virkning forandres.

Betingelsen er således både intuitivt appellerende, logisk og ontologisk meningsfuld. Det argumenteres der ikke imod. Argumentet er, at der i samfundsvidenskaben er praktiske problemer forbundet med at bruge tidsrækkefølge til at identificere størrelsen af den kausale effekt af x på y .

Således er betingelsen i praksis irrelevant, hvis der er stiafhængighed i det empiriske system, der undersøges. Stiafhængighed betyder løst formuleret, at fortidige variabelværdier påvirker nutidige variabelværdier. For den kausale identifikation implicerer stiafhængighed, at hvis x på et givet tidspunkt er endogen, så vil x også være endogen på et tidligere tidspunkt. Dette er nogenlunde intuitivt, for en sammenhæng mellem samme variabel på to tidspunkter (altså stiafhængighed) betyder, at eventuelle endogenitetsproblemer blot flyttes tilbage i tid, hvis vi erstatter en nutidig variabelværdi med en tidligere: Sammenhængen mellem x_{t_0} og x_{t-1} flytter blot endogenitetsproblemet fra t_0 til t_{-1} , hvis vi erstatter x_{t_0} med x_{t-1} i en empirisk analyse. Dette gælder uanset, om det

er den afhængige variabel, der er stiafhængig, den uafhængige eller ”alt det andet, som korrelerer med x eller y ” (den interesserede læser henvises til noten for den formelle udledning).⁶

Antag til eksempel at x i figur 1 er demokrati og y økonomisk udvikling. Her er x endogen, idet økonomisk udvikling virker tilbage på demokrati (Przeworski et al. 2000: 90), ligesom der er et komplekst samspil af vanskeligt observerbare tredjevariable, som varierer med både x og y (se evt. Acemoglu et al., 2008). Hjælper det i denne situation at betragte sammenhængen mellem demokrati tilbage i tid og økonomisk udvikling i nutiden, når vi skal identificere b_y ? Nej – og det skyldes den massive stiafhængighed i systemet. Således er både demokrati og økonomisk udvikling træge variable, der afhænger af sig selv i fortiden. Det betyder, at enhver endogenitet tilbage i tid også skaber nutidig endogenitet: F_x vil demokrati tilbage i tid (på t_{-j}) afhænge (positivt) af økonomisk udvikling på t_{-j} . Hvis økonomisk udvikling i dag afhænger positivt af sig selv i fortiden, vil sammenhængen mellem demokrati i fortiden og økonomisk udvikling i dag indeholde den træge (positive) effekt af økonomisk udvikling i fortiden og dermed overvurdere den kausale effekt af demokrati på udvikling. Eksemplet illustrerer, og note *vi* viser, hvad der menes med, at kravet om tidsrækkefølge i empiriske undersøgelser med stiafhængighed er irrelevant: Hvis x_{t_0} er endogen, vil x_{t-1} også være det.

Den praktiske anvendelighed af kriteriet vanskeliggøres yderligere af strategisk handlen. Som Asbjørn Sonne Nørgaard (2007: 239) rigtigt har understreget, antages det i megen samfundsvidenskab, at aktører handler på baggrund af deres forventninger til fremtiden. Det er fx en af de bærende antagelser i spilteorien, hvor spil over flere perioder som bekendt løses ”bagfra”. Den rationelle aktør begynder så langt ude i fremtiden, som hendes tidshorisont tillader, træffer det bedst mulige valg i denne fremtidige periode og regner sig derefter baglæns tilbage til nutiden. Her er det naturligvis ikke fremtiden, der påvirker nutiden – det er den nutidige *forventning* om, hvad der sker på et senere tidspunkt, der er årsag til agenternes handling i nutiden. Dette skaber problemer i de empiriske analyser, hvor agenternes forventninger ikke observeres: Der eksisterer simpelthen ikke en observerbar årsag på t_0 for udfaldet på t_1 .

At dette kan have praktiske konsekvenser illustrerer Adam Przeworski et al.’s demokratiindeks (2000: 23-28): Betragt et spil over to perioder. I første periode beslutter en regering, om den vil afholde valg; i den anden beslutter vælgerne, hvorvidt regeringen skal fortsætte. Regeringens forventning til valgresultatet i den anden periode påvirker beslutningen i første periode således, at regeringen kun afholder valg, når den har en forventning om at blive genvalgt. Forfatterne lægger dette simple spil til grund og argumenterer for, at selvom

valghandlingen på t_0 ideelt set er årsag til det folkelige mandat på t_1 (2000: 27), så kan det ikke afvises, at valget kun blev afholdt, fordi regeringen forventede at vinde – og ikke ville være blevet afholdt, hvis regeringen havde forventet at tabe. Problemet er, at forfatterne ikke kan observere regeringens forventninger *ex ante*, kun udfaldene (valg, ikke-valg; valgsejr, valgnederlag). Så på grund af strategisk handlen og ikke-observerbare forventninger vælger Przeworski et al. at klassificere de lande, hvori regeringen afholder valg, men aldrig taber, som diktaturer. Eksemplet illustrerer, hvordan strategisk handlen kan betyde, at udfald (valget på t_0) kan have ikke-observerbare årsager, samt at forfatterne i mangel af bedre søger årsagen til udfaldet på t_0 i udfaldet på t_1 (valgresultatet). Det illustrerer vanskeligheden ved kravet om tidsrækkefølge i situationer med strategisk handlen.

Det skal understreges, at der ikke argumenteres imod tidsrækkefølgens ontologiske gyldighed i kausale betragtninger. Det er formentlig ontologisk uangribeligt, at årsag indtræder før virkning. Det påpeges, at kravet kan have nogle svagheder i praktisk empirisk forskning: Hvis der er stiafhængighed i det empiriske system, der undersøges, løser kravet ikke eventuelle endogenitetsproblemer men rykker dem blot tilbage i tiden. Hvis der er strategisk handlen, kan den teoretisk relevante årsag til et givet udfald ofte ikke observeres.

Teoretisk forklaring og kausal identifikation

Ved en forklaring forstår vi, at den mekanisme, der fører fra den uafhængige til den afhængige variabel, specificeres (Elster, 2007: 21; Gerring, 2010: 1502). Ved en teoretisk forklaring forstår vi, at mekanismen sandsynliggøres gennem teoretiske argumenter men ikke nødvendigvis identificeres empirisk. Der argumenteres i dette afsnit for, at kriteriet nok bestyrker os i, at den kausale effekt er en delmængde af den observerede korrelation, men at kriteriet ikke hjælper os til at identificere den kausale effekt. Desuden argumenteres der for, at en teoretisk forklaring ikke er nødvendig for at identificere en kausal effekt.

Betragt igen systemet i figur 1 og lad fortsat x repræsentere demokrati og y økonomisk udvikling. Der kan da relativt let gives teoretiske forklaringer på såvel effekten fra x til y (b_y) som det modsatte (b_x). I dette system indeholder den observerede korrelation mellem x og y altså størrelsen ($b_y + b_x + \text{øvrige sammenhænge}$). At b_y kan begrundes teoretisk, udelukker dog ikke en teoretisk begrundelse af b_x (eller af øvrige sammenhænge) og retfærdiggør dermed ikke, at korrelationen ($b_y + b_x + \text{øvrige sammenhænge}$) fortolkes som b_y . Mere generelt kan man sige, at den kausale relation, vi interesserer os for (b_y), er en delmængde af den observerede korrelation ($b_y + b_x + \text{øvrige sammenhænge}$). Da formålet med at opstille kriterier for kausal identifikation netop er at udsønde

denne delmængde, hjælper kriteriet om teoretisk forklaring ikke til at udsondre b_y , hvis der også kan gives teoretiske forklaringer på de øvrige sammenhænge i systemet (hvad der ofte kan i samfundsvidenskaben). Kriteriet kan i stedet ses som et minimumskriterium, der hjælper os til at udsondre nonsens-korrelationer: Der kan ikke gives en teoretisk tilfredsstillende forklaring på, at antallet af storke skulle påvirke antallet af børnefødsler, og korrelationen kan derfor afskrives. Kravet om en teoretisk forklaring er således et minimumskriterium, der kan bestyrke os i, at b_y ikke er nonsens og således faktisk er en delmængde af korrelationen. Men det hjælper os ikke til at forkaste b_x eller de øvrige sammenhænge.

Kravet om teoretisk forklaring er heller ikke nødvendigt for at identificere b_y : Man kan godt identificere en kausal sammenhæng uden at kunne forklare den. Betragt følgende generelle eksempel: I et eksperimentelt forskningsdesign placeres en gruppe forsøgspersoner efter randomiseret udvælgelse i to grupper, en behandlingsgruppe og en kontrolgruppe. Behandlingsgruppen modtager en given stimulans, x , som kontrolgruppen ikke modtager, hvorefter begge grupper måles med hensyn til et relevant udfald, y . I dette forskningsdesign kan den gennemsnitlige kausale påvirkning af x på y identificeres som den gennemsnitlige forskel i udfaldet i behandlings- og kontrolgruppen. Det eksperimentelle forskningsdesign identificerer dermed, hvor meget x i gennemsnit påvirker y , men det giver ikke i sig selv en kausal forklaring på, hvorfor det sker.

Som et konkret eksempel kan vi betragte Ignaz Semmelweis' opdagelse af sammenhængen mellem lægers håndvask og reduceret dødelighed i barselssengen. Resultatet blev afskrevet af samtiden – Louis Pasteur havde nemlig endnu ikke opdaget bakterier, og dermed var der endnu ikke etableret et teoretisk fundament for Semmelweis' opdagelse (Best og Neuhammer, 2004: 233). At sammenhængen, da den blev opdaget, ikke kunne forklares teoretisk, betød imidlertid ikke, at den ikke var kausal. Eksemplerne illustrer altså, at en teoretisk forklaring ikke er en nødvendig betingelse for at identificere størrelsen af en kausal sammenhæng.

Det skal understreges, at dette ikke er et argument imod anvendelsen af teori i empiriske undersøgelser. Således illustrerer eksemplet med håndvask og dødelighed netop, hvordan samspillet mellem empiri og teori driver forskningen fremad. Opdagelsen af bakterier gjorde det muligt at færdiggøre den kausale kæde fra håndvask til reduceret dødelighed og dermed forklare Semmelweis' opdagelse. Teori er også vigtigt i praktisk empirisk arbejde, der ikke blot søger kausal identifikation (størrelsen af b_y) men kausal forklaring (mekanismen der forbinder x og y). Hvis man fx vil forklare, hvorfor beskyttelse af ejendomsretten på langt sigt påvirker den økonomiske udvikling, er det ikke

nok at henvise til størrelsen af b_j . En forklaring forudsætter, at det specificeres, om påvirkningen går gennem investeringsniveau, teknologiudvikling eller noget andet. Forklaringen forudsætter teoretiske antagelser, for ved at sige at veldefinerede ejendomsrettigheder fremmer investeringslysten, gøres teoretiske antagelser om aktørpræferencer.

Teori spiller således en meget væsentlig rolle i empiriske analyser. Artiklens ærinde er dog ikke, hvad der driver forskningen fremad, eller hvordan empiriske sammenhænge forklares. Artiklens ærinde er så præcist som muligt at få indsættet, hvilke kriterier der kan stilles til en observeret sammenhæng, før vi kan sige at have identificeret størrelsen af den kausale effekt. I denne snævre sammenhæng er en teoretisk forklaring formelt set hverken nødvendig eller tilstrækkelig.

Kausal identifikation i observationsstudier: nøglen er design

I dette afsnit bevæger vi os fra det abstrakte til det praktiske og eksemplificerer, hvordan empirisk samfundsforskning i praksis kan søge at opnå kausal identifikation – selv under det vilkår at det meste hænger sammen, og ingen variabel dermed umiddelbart er eksogen. Først skal der drages en parallel mellem den beskrevne tilgang til kausal identifikation og det eksperimentelle forskningsdesign.

Som nævnt indledningsvis er en af de fundamentale vanskeligheder ved kausal inferens, at vi ikke kan observere, hvad der ville have været virkningen, hvis årsagen (kontrafaktisk) var udeblevet. Vi kan fx ikke spole tiden tilbage til den hypotetiske situation før en eksplosion og observere, hvad der ville være sket, hvis gnisten var udeblevet. Derfor forsøger mange typer forskningsdesign at konstruere et tillempet kontrafakta ved hjælp af observerbart data. Dette tillempede kontrafakta kan så sammenlignes med det faktisk forekommende og derved danne baggrund for en kausal slutning. Dette er udgangspunktet i den såkaldte Rubins kausalmodel (Winship og Sobel, 2001), men betingelsen for at kunne opnå kausal inferens er også i denne model, at den uafhængige variabel opfylder (a) og (b).

Logikken bag det tillempede kontrafakta kommer klarest til udtryk i et eksperimentelt forskningsdesign. Her sikrer den tilfældige udvælgelse af forsøgspersoner, at behandlingsgruppen og kontrolgruppen gennemsnitligt set er ens på alle dimensioner bortset fra, at den ene gruppe får et stimuli, som den anden ikke får. Kontrolgruppen fungerer dermed som et gennemsnitligt kontrafakta for behandlingsgruppen, og derfor kan vi drage en valid kausal slutning om den gennemsnitlige kausale virkning af det stimuli, behandlingsgruppen udsættes for.

Parallellen er her, at den tilfældige udvælgelse netop sikrer, at stimuli (den uafhængige variabel) er eksogen og dermed opfylder (a) og (b). Hvis x opfylder (a) og (b), er værdien af x i gennemsnit den eneste (relevante) forskel på de undersøgte observationer. Der er således ingen bagvedliggende faktorer, der forklarer x , som også påvirker de øvrige variable, ligesom ingen af de øvrige variable påvirker x . At opfylde (a) og (b) betyder derfor, at den uafhængige variabel i praksis er tilfældigt fordelt på tværs af observationerne, og efter samme logik som i det eksperimentelle forskningsdesign betyder eksogeniteten, at der kan drages en valid kausal slutning om en gennemsnitlig årsagssammenhæng.

Imidlertid minder samfundet mere om figur 1 end om et eksperimentelt laboratorium: Næsten alle faktorer påvirker hinanden, og det er ofte vanskeligt at hævde, at en relevant samfundsmæssig variabel i en relevant samfundsmæssig sammenhæng er eksogen eller tilfældigt fordelt. Dermed er det også vanskeligt at identificere kausale sammenhænge i empirisk samfundsforskning.

Den konstruktive lære er imidlertid, at selvom det er vanskeligt, er det ikke umuligt. Det kræver dog et velgennemtænkt forskningsdesign. Da artiklen her ikke er tænkt som et katalog over empiriske design, nøjes jeg med at illustrere dette med et eksempel på, hvordan man kan designe observationsstudier således, at den uafhængige variabel opfylder (a) og (b).

Antag for eksempel, at vi ønsker at identificere den kausale påvirkning af ændringer i national indkomst på sandsynligheden for borgerkrig på tværs af lande i Afrika. Dette er en vanskelig empirisk problemstilling, idet hverken (a) eller (b) er opfyldt. For det første er der omvendt kausalitet, idet sandsynligheden for borgerkrig påvirker den nationale indkomst, lige såvel som den nationale indkomst påvirker sandsynligheden for borgerkrig. Dette problem kan ikke løses ved simpelthen at betragte værdier for national indkomst på et tidligere tidspunkt end borgerkrigen: Den nationale indkomst bliver skabt af borgere, som træffer deres beslutning om at investere, opfinde eller uddanne sig på baggrund af deres forventninger om fremtiden – herunder naturligvis deres forventning om sandsynligheden for borgerkrig. Desuden er der betydelig stiftafhængighed i både national indkomst og sandsynligheden for borgerkrig, så som vist ovenfor vil det ikke fjerne endogenitetsproblemet at benytte tidligere værdier for national indkomst. For det andet er kriterium (b) heller ikke opfyldt, idet der næsten med sikkerhed er historiske og kulturelle faktorer, som ikke kan observeres, men som hænger sammen med både den nationale indkomst og sandsynligheden for borgerkrig. Alt dette betyder, at den nationale indkomst ikke er eksogen i denne sammenhæng, og størrelsen af den kausale påvirkning herfra på sandsynligheden for borgerkrig kan ikke umiddelbart identificeres.

Edward Miguel et al. (2001) løste problemet ved i et kendt studie at benytte variation i nedbøren som kilde til eksogen variation i den nationale indkomst i afrikanske lande. Da landbrug stadig spiller en betydelig økonomisk rolle i Afrika, har mængden af nedbør målbare nationaløkonomiske konsekvenser. Derfor benyttede forfatterne kun den del af den nationale indkomst, som afhænger af nedbøren, til at identificere den kausale påvirkning af indkomst på sandsynligheden for borgerkrig. Dette giver kausal identifikation, fordi nedbør i denne sammenhæng opfylder (a) og (b): Sandsynligheden for borgerkrig påvirker ikke mængden af nedbøren, ligesom det nok er begrænset, i hvilken grad at de faktorer, som påvirker nedbøren (fx atmosfæriske trykforhold), påvirker sandsynligheden for borgerkrig andet end gennem den nationale indkomst.

I eksemplet benyttes nedbør formelt set som instrument-variabel for national indkomst. Pointen her er dog ikke det enkelte forskningsdesign – dertil er det alt for rudimentært beskrevet. Pointen er, at selv under det vilkår, at det meste hænger sammen og kun få variable kan siges at være eksogene i udgangspunktet, så findes der standardiserede forskningsdesign som er skræddersyet til kausal identifikation (jf. henvisningerne i note 1).

Konklusion

Lidt stiliseret kan man opstille to idealtypiske syn på kausal identifikation i observationsstudier. Den ene idealtipe kan repræsenteres ved en af de første multivariate analyser i samfundsvidenskaben, Yules analyse af årsager til fattigdom i England fra 1899. Yule fandt samvariation, men sprogbrugen var kausal, og spørgsmålet om kausal inferens blev gemt i en berømt fodnote: "Strictly, for *due to* read *associated with*" (1899: 270, note 25). Den anden idealtypiske position siger, at kun kontrollerede eksperimenter – aldrig observationsstudier – kan give anledning til kausale slutninger. Argumentet i artiklen her er tænkt som en mellemposition. Det er ikke nok at lade som om, at korrelation er kausalitet og feje indvendinger ind under gulvtæppet og ned i fodnoter. Det bidrager selvsagt ikke til fagets praktiske relevans. Det bidrager dog lige så lidt at ignorere virkeligheden uden for det eksperimentelle laboratorium, en virkelighed som nu engang må studeres gennem observationsstudier. Den pragmatiske mellemposition siger, at selvom vi i mange situationer ikke kan gøre det perfekte, må vi i det mindste gøre det, så godt vi kan. Et første skridt er her at få afklaret, hvilke kriterier der kan stilles, før vi har identificeret en kausal effekt.

I definitionen af en årsagssammenhæng blev asymmetrien understreget: Årsag har kausal prioritet og forårsager virkning; det modsatte er ikke tilfældet. Det blev demonstreret grafisk og matematisk, at den asymmetri, der tillader identifikation af størrelsen af den kausale effekt af årsag på virkning, er ek-

sogenitet. Jeg argumenterede i den forbindelse for at de tre gængse kriterier (fravær af tredje-variabel, tidsrækkefølge, teori) i mange praktiske anvendelser kan være upræcise. Det er mere stringent at argumentere for, at (a) der ikke må være omvendt kausalitet, samt (b) at intet af det, som korrelerer med den uafhængige variabel, må korrelere med noget, som korrelerer med den afhængige.

Argumentet er ikke, at disse kriterier omdanner vores rodede virkelighed til et eksperimentelt laboratorium og gør statskundskab til naturvidenskab. Argumentet er, at vi så klart som muligt formulerer de endogenitetsproblemer, der præger statskundskabens emnefelt, at vi så klart som muligt forsøger at tage hånd om dem, og håbet er, at vi med tiden begynder at uddanne vore studerende i de metoder, som er skræddersyet til kausal identifikation – også efter det krævende kriterium, artiklen her har skitseret.

Noter

1. Forfatteren takker Politicas anonyme bedømmere samt Mogens Kamp Justesen, Sigge Winther Nielsen, Asmus Leth Olsen og Asger Moll Wingender for nyttige kommentarer.
2. Læseren henvises til udmærkede engelsksprogede bidrag, som “katalogiserer” metoder til kausal inferens. Herunder Winship og Morgan (1999); Blundell og Costa Dias (2008).
3. En inus-årsag er ikke i sig selv hverken nødvendig eller tilstrækkelig men en del af et sæt af årsager, som kollektivt frembringer virkningen (og er kollektivt tilstrækkelige). Virkningen kunne dog have været frembragt af andet (sættet af årsager er ikke kollektivt nødvendigt).
4. Den probabilistiske forståelse af kausalitet er symmetrisk, fordi $p(y|x) > p(y)$ implicerer, at $p(x|y) > p(x)$. Dette følger af Bayes' lov: $p(y|x) \cdot p(x) = p(x|y) \cdot p(y)$. Hvis udtrykket for $p(y|x) [= p(x|y) \cdot p(y) / p(x)]$ substitueres ind i den første ulighed, kan udtrykket let manipuleres, således at det omvendte resultat følger (nemlig at y er kausalt relevant for x).
5. Ved at gange (1') og (2'') med hinanden og tage den forventede værdi fås $E[xy] = b_y E[x^2] - E[u_x u_y]$. Idet $E[u_x u_y]$ er nul, kan udtrykket omskrives til den velkendte OLS-estimator (for den simple lineære regression uden skæring): $b_y = E[xy] / E[x^2]$.
6. Antag, at det sidste gælder – at u_y er stiafhængig og dermed afhænger af sig selv i fortiden som defineret af Scott Page (2006): $u_{y,t} = \alpha u_{y,t-1} + e_t$. α måler graden af persistens og e_t repræsenterer tilfældige stød med konstant varians og middelværdi 0. Antag at vi i denne situation vil estimere en sammenhæng mellem to kontinuerte variable, x_t og y_t , hvor x_t betragtes som årsag til y_t , svarende til modellen $y_t = a + b x_t + u_{y,t}$. Antag imidlertid, at (b) ikke er opfyldt, således at $E[x_t u_{y,t}] \neq 0$. I denne situation løser det intet at erstatte den uafhængige variabel på tidspunkt t med samme va-

riabel på et tidligere tidspunkt: Selvom den uafhængige variabel kommer før den afhængige variabel, er (b) stadig ikke opfyldt, fordi u_y udviser stiafhængighed. Det indses let ved substitution: $E[x_{t-1}u_y] = E[x_{t-1}(\varrho^*u_{y,t-1} + e_y)] = \varrho^*E[x_{t-1}u_{y,t-1}] \neq 0$. Altså er det kun i de sjældne tilfælde, hvor der ingen stiafhængighed er (svarende til $\varrho=0$), at det løser endogenitetsproblemet at erstatte x_t med x_{t-1} .

Litteratur

- Acemoglu, Daron, Simon Johnson, James Robinson og Pierre Yared (2008). Income and Democracy. *The American Economic Review* 98 (3): 808-842.
- Andersen, Lotte Bøgh (2010). Forskningskriterier, pp. 97-113 i Lotte Bøgh Andersen, Robert Klemmensen og Kasper Møller-Hansen (red.), *Metoder i statskundskab*. København: Hans Reitzels Forlag.
- Best, M. og D. Neuhaser (2004). Ignaz Semmelweis and the Birth of Infection Control. *Quality Health Care* 13: 233-234.
- Blundell, Richard og Monica Costa Dias (2008). *Alternative Approaches to Evaluation in Empirical Microeconomics*. <http://www.ucl.ac.uk/~uctp39a/Blundell-CostaDias-27-09-2008.pdf>.
- Cartwright, Nancy (2007). *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge, UK: Cambridge University Press.
- Elster, Jon (2007). *Explaining Social Behavior. More Nuts and Bolts for the Social Sciences*. Cambridge, UK: Cambridge University Press.
- Hoover, Kevin (2001). *Causality in Macroeconomics*. Cambridge, UK: Cambridge University Press.
- Keuzenkamp, Hugo (2000). *Probability, Econometrics, Truth: The Methodology of Econometrics*. Cambridge, UK: Cambridge University Press.
- Mackie, John. (1974). *The Cement of the Universe: A Study of Causation*. Oxford: Clarendon Library of Logic and Philosophy.
- Miguel, Edward, Shanker Satyanath og Ernest Sergenti (2001). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy* 112 (4): 725-753.
- Nørgaard, Asbjørn Sonne (2007). God statskundskab: Heksekunst eller håndværk. *Politica* 39 (3): 233-255.
- Page, Scott (2006). Path Dependence. *Quarterly Journal of Political Science* 1: 87-115.
- Przeworski, Adam, Michael M. Alvarez, José Antonio Cheibub og Fernando Limongi (2000). *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*. Cambridge: Cambridge University Press.
- Psillos, Stathis (2009). Regularity Theories, pp. 131-157 i Helen Beebe, Christopher Hitchcock og Peter Menzies (red.), *Oxford Handbook of Causation*. Oxford: Oxford University Press: 131-157.

- Russell, Bertrand (1913). On the Notion of Cause. *Proceedings of the Aristotelian Society 1912-1913*: 1-26.
- Simon, Herbert (1952). On the Definition of the Causal Relation. *The Journal of Philosophy* XLIX (16): 517-528.
- Simon, Herbert (1953). Causal Ordering and Identifiability, pp. 49-75 i William C. Hood og T.C. Koopmans (red.), *Studies in Econometric Method. Cowles Commission for Research in Economics. Monograph No. 14*. New York: John Wiley and Sons Inc.
- Simon, Herbert (1954). Spurious Correlation: A Causal Interpretation. *Journal of the American Statistical Association* 49: 467-479.
- Sløk, Johannes, Mogens Pihl og Erik Lund (2007). *De europæiske ideers historie*. København: Gyldendal.
- Suppes, Patrick (1970). A Probabilistic Theory of Causality. *Acta Philosophica Fennica* XXIV.
- Yule, G. Udny (1899). An investigation into the causes of changes in pauperism in England, chiefly during the last two intercensal decades. *J. Roy. Statist. Soc.* 62: 249-295.
- Winship, Christopher og Stephen Morgan (1999). The Estimation of Causal Effects From Observational Data. *Annual Review of Sociology* 25: 659-707.
- Winship, Christopher og Michael Sobel (2001). *Causal Inference in Sociological Studies*. http://www.wjh.harvard.edu/soc/faculty/winship/cfa_papers/causalinference.pdf.